

Pengenalan Sains Komputasi - Fitting data

Ridlo W. Wibowo || 20912009

October 31, 2012

Problem.

Fitting 2 set data yang ada dengan menggunakan software atau tools fitting dari bahasa pemrograman anda.

Documentation.

Data yang dipilih dari sumber:

1. NorrisEM.txt (untuk coba-coba terlebih dahulu)
2. FilipEM.txt
3. Gauss1.txt

Software yang digunakan adalah **gnuplot** (OS: Ubuntu).

Setelah dilakukan perbaikan format data (karena tidak bisa dibaca oleh **gnuplot - ubuntu**), maka proses fitting dilakukan dengan cara sederhana, yaitu menentukan fungsi fitting yang ingin kita gunakan lalu menggunakan fungsi **fit** yang ada di dalam **gnuplot** (lihat dokumentasi **gnuplot**). Fungsi ini adalah implementasi algoritma *Levenberg-Marquardt* untuk nonlinear least-squares (NLLS) pada **gnuplot**.

1. **NorrisEM.txt**

Kita ingin melakukan fitting data ini menggunakan fungsi linier (garis lurus) dengan persamaan:

$$y = a + bx \quad (0.1)$$

dengan fungsi *fit* yang ada pada **gnuplot** kita ingin menentukan konstanta *a* dan *b* yang menjadikan fungsi di atas dapat menggambarkan data yang ada, atau sering disebut regresi linier.

Langkah-langkahnya, pertama kita buat file input untuk gnuplot agar pengerjaan lebih mudah (*norris.in*):

```
1 set output "norris.png"
2 set xlabel "x"
3 set ylabel "y = f(x)"
4 f(x) = a + b*x
5 fit f(x) "Norris.txt" using 1:2 via a,b
6 plot "Norris.txt" u 1:2 title "data", f(x) title "fitting function"
```

kemudian dapat dijalankan dengan perintah sederhana (di ubuntu):

```
$ gnuplot < norris.in
```

gnuplot akan melakukan iterasi dan menghasilkan fitting terbaik dengan toleransi tertentu, dalam kasus ini:

```
After 5 iterations the fit converged.
final sum of squares of residuals : 26.5049
rel. change during last iteration : -7.49057e-06
```

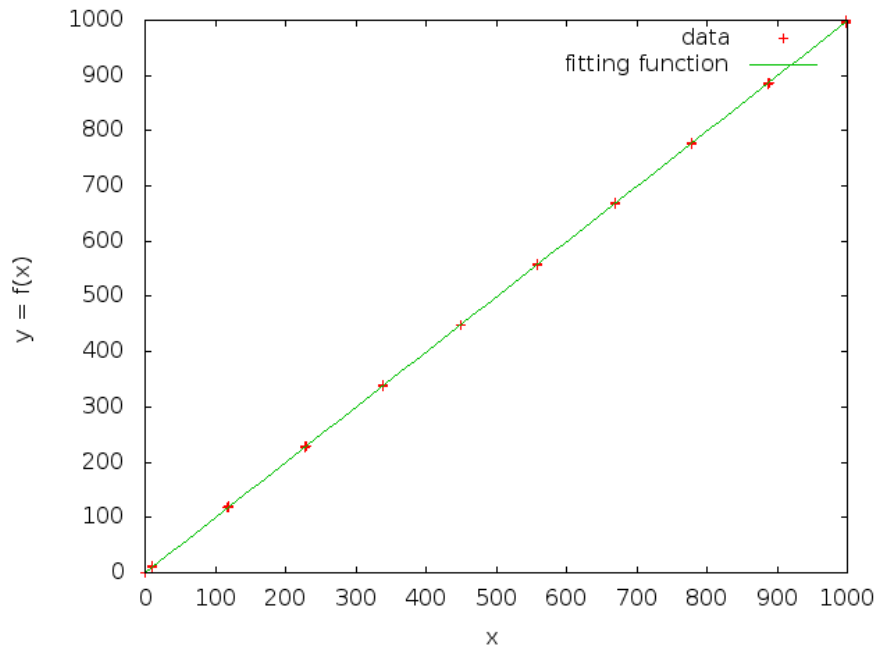
```
degrees of freedom      (FIT_NDF)                : 34
rms of residuals        (FIT_STDFIT) = sqrt(WSSR/ndf) : 0.882925
variance of residuals (reduced chisquare) = WSSR/ndf  : 0.779556
```

Final set of parameters	Asymptotic Standard Error
=====	=====
a = 0.264389	+/- 0.2322 (87.84%)
b = 0.997881	+/- 0.000428 (0.04289%)

correlation matrix of the fit parameters:

	a	b
a	1.000	
b	-0.774	1.000

dan menggunakan perintah di atas kita juga akan peroleh plot data dan fungsi hasil fittingnya:



NorrisEM.txt dan fungsi fittingnya.

diperoleh fungsi fitting:

$$y = 0.264389 + 0.997881x \quad (0.2)$$

2. FilipEM.txt

Kita ingin melakukan fitting data ini menggunakan fungsi polinom dengan persamaan:

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 + a_6x^6 + a_7x^7 + a_8x^8 + a_9x^9 + a_{10}x^{10} \quad (0.3)$$

dengan fungsi *fit* yang ada pada gnuplot kita ingin menentukan 11 konstanta pada persamaan di atas. Langkahnya sama seperti sebelumnya, pertama kita buat file input untuk gnuplot (*filip.in*):

```
set output "filip.png"
set xlabel "x"
set ylabel "y = f(x)"
f(x) = a0 + a1*x + a2*x**2 + a3*x**3 + a4*x**4 + a5*x**5 + a6*x**6 +
      a7*x**7 + a8*x**8 + a9*x**9 + a10*x**10
fit f(x) "Filip.txt" using 1:2 via a0,a1,a2,a3,a4,a5,a6,a7,a8,a9,a10
plot "Filip.txt" using 1:2 title "data", f(x) title "fitting
      polynomial"
```

kemudian jalankan gnuplot:

```
$ gnuplot < filip.in
```

Hasil:

After 16 iterations the fit converged.

final sum of squares of residuals : 2.88774

rel. change during last iteration : -9.30383e-06

degrees of freedom (FIT_NDF) : 71
rms of residuals (FIT_STDFIT) = sqrt(WSSR/ndf) : 0.201674
variance of residuals (reduced chisquare) = WSSR/ndf : 0.0406724

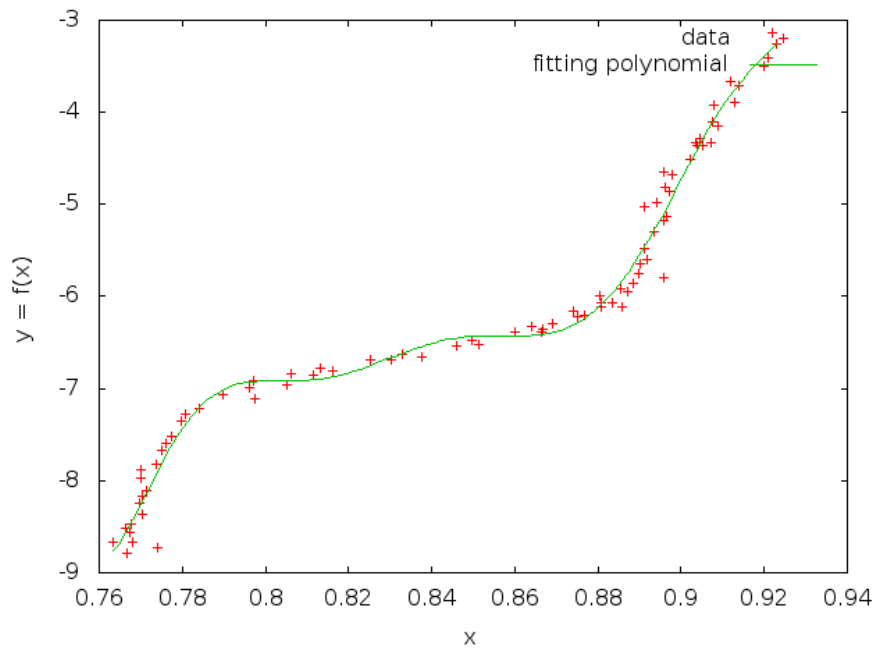
Final set of parameters	Asymptotic Standard Error
=====	=====

a0	= 2.00527e+07	+/- 2.549e+08	(1271%)
a1	= 1.0012e+09	+/- 2.59e+09	(258.7%)
a2	= -9.27085e+09	+/- 1.495e+10	(161.3%)
a3	= 3.31814e+10	+/- 5.965e+10	(179.8%)
a4	= -5.90527e+10	+/- 1.597e+11	(270.4%)
a5	= 4.47287e+10	+/- 2.865e+11	(640.6%)
a6	= 1.95601e+10	+/- 3.478e+11	(1778%)
a7	= -7.31835e+10	+/- 2.831e+11	(386.9%)
a8	= 6.65244e+10	+/- 1.486e+11	(223.4%)
a9	= -2.83802e+10	+/- 4.561e+10	(160.7%)
a10	= 4.87144e+09	+/- 6.229e+09	(127.9%)

correlation matrix of the fit parameters:

	a0	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10
a0	1.000										
a1	-0.887	1.000									
a2	0.550	-0.872	1.000								
a3	-0.212	0.634	-0.930	1.000							
a4	-0.015	-0.435	0.816	-0.970	1.000						
a5	0.160	0.290	-0.712	0.917	-0.986	1.000					
a6	-0.255	-0.184	0.624	-0.861	0.957	-0.992	1.000				
a7	0.319	0.103	-0.549	0.807	-0.923	0.973	-0.995	1.000			
a8	-0.363	-0.039	0.485	-0.756	0.886	-0.949	0.981	-0.996	1.000		
a9	0.394	-0.012	-0.428	0.707	-0.848	0.922	-0.963	0.986	-0.997	1.000	
a10	-0.415	0.054	0.377	-0.661	0.811	-0.893	0.941	-0.971	0.988	-0.997	1.000

plot data dan fungsi hasil fittingnya:



FilipEM.txt dan fungsi fittingnya.

3. Gauss1EM.txt

Kita ingin melakukan fitting data ini menggunakan fungsi gauss dengan persamaan:

$$y = f(x) = a_0 e^{-a_1 x} + a_2 e^{\frac{-(x-a_3)^2}{a_4^2}} + a_5 e^{\frac{-(x-a_6)^2}{a_7^2}} \quad (0.4)$$

dengan fungsi *fit* yang ada pada gnuplot kita ingin menentukan 7 konstanta pada persamaan di atas.

Data yang ada ternyata harus ditukar kolomnya apabila kita ingin melakukan fitting suatu fungsi.

Langkahnya hampir sama seperti sebelumnya, namun untuk kasus ini setelah dicoba ternyata tidak akan konvergen apabila kita tidak memasukkan tebakan awal nilai konstanta yang cukup dekat dengan data yang ada. Setelah memperkirakan berapa nilai untuk tebakan awal sesuai data yang ada, kita dapat memasukkannya ke dalam file input untuk gnuplot (*gauss.in*) sehingga menjadi seperti berikut ini:

```
set term png
set output "gauss.png"
set xlabel "x"
set ylabel "y = f(x)"
a0 = 80
```

```

a1 = 0.01
a2 = 100
a3 = 50
a4 = 15
a5 = 70
a6 = 170
a7 = 18
f(x) = a0*exp(-a1*x) + a2*exp((- (x-a3)**2)/(a4**2)) + a5*exp((- (x-a6)
**2)/(a7**2))
fit f(x) "Gauss.txt" using 2:1 via a0,a1,a2,a3,a4,a5,a6,a7
plot "Gauss.txt" using 2:1 title "data", f(x) title "fitting function"

```

kemudian jalankan gnuplot:

```
$ gnuplot < gauss.in
```

Hasil:

After 8 iterations the fit converged.

final sum of squares of residuals : 1315.82

rel. change during last iteration : -2.38286e-10

```

degrees of freedom      (FIT_NDF)                : 242
rms of residuals        (FIT_STDFIT) = sqrt(WSSR/ndf)    : 2.3318
variance of residuals (reduced chisquare) = WSSR/ndf    : 5.43728

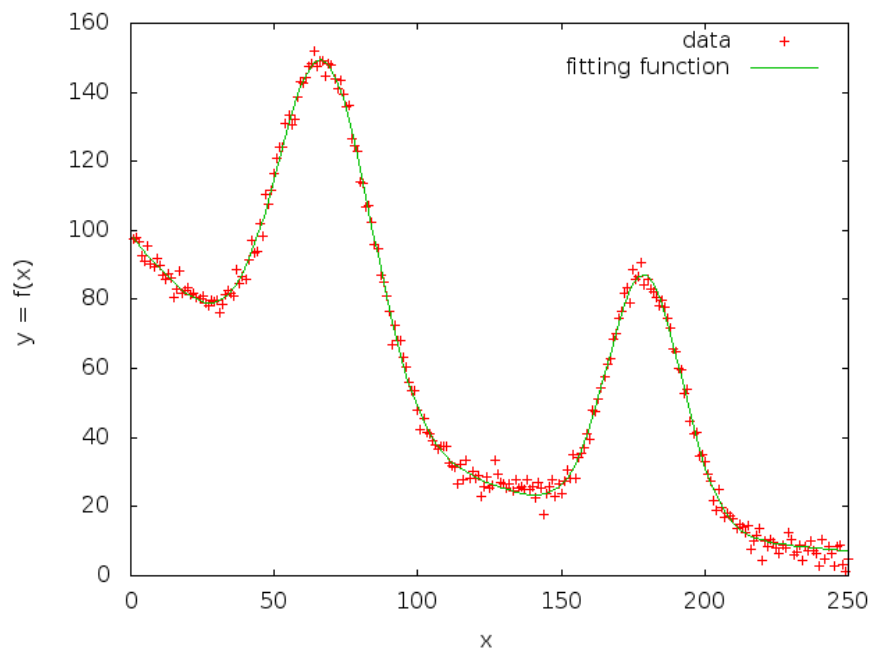
```

Final set of parameters		Asymptotic Standard Error	
=====		=====	
a0	= 98.778	+/- 0.5754	(0.5825%)
a1	= 0.0104972	+/- 0.0001142	(1.088%)
a2	= 100.49	+/- 0.5883	(0.5855%)
a3	= 67.4811	+/- 0.1046	(0.155%)
a4	= 23.1297	+/- 0.1745	(0.7544%)
a5	= 71.9943	+/- 0.626	(0.8696%)
a6	= 178.998	+/- 0.1244	(0.06948%)
a7	= 18.3893	+/- 0.2014	(1.095%)

correlation matrix of the fit parameters:

	a0	a1	a2	a3	a4	a5	a6	a7
a0	1.000							
a1	0.494	1.000						
a2	-0.084	0.328	1.000					
a3	0.281	0.237	0.026	1.000				
a4	-0.271	0.321	-0.175	-0.019	1.000			
a5	0.076	0.301	0.122	0.056	0.138	1.000		
a6	0.010	-0.055	-0.030	-0.005	-0.039	-0.013	1.000	
a7	0.107	0.470	0.195	0.084	0.222	-0.324	-0.038	1.000

plot data dan fungsi hasil fittingnya:



Gauss1EM.txt dan fungsi fittingnya.

== *** ==